

REMARKS

Reconsideration and further examination of the subject patent application in light of the present Amendment and Remarks is respectfully requested. Claims 1 and 8-14 are currently pending in the application and stand rejected.

Independent claims 1 and 8 have been amended to point out that referencing document references or refers to a referenced document by means of links, and to highlight that the aggregated document created does not exist as an entity in the collection of documents searched by a user, meaning one cannot find such an aggregated document on the Internet or other network. The claims as amended are supported in the specification. For example, with reference to Figure 2, documents 30, 32 and 36, or documents 30, 32, 34, 36, 38, 40 and 42 are deemed to be a single aggregate or logical document (Specification, page 8, 4th par.).

Rejection under 35 U.S.C. §102

Claims 1, 10, and 14 stand rejected under 35 U.S.C. §102(e) as being anticipated by Snyder (US 6,643,641). Applicant respectfully traverses the rejection. The Examiner cites applicant's claimed steps of aggregating, indexing, searching, and returning, and then cites various sections of Snyder as disclosing this subject matter. Applicant disagrees that Snyder describes this subject matter, as discussed below:

Aggregation and Indexing

The Examiner cites at least four sections in Snyder as describing applicant's claimed subject matter. First, the Examiner cites Snyder, Col. 9, lines 57-61. This section describes typical data included in web pages, which are stored in the buffer. There is no mention whatsoever of any aggregation process or aggregated documents. Applicant does not understand

the relevance of this section and respectfully requests that the Examiner provide an explanation of how this section describes applicant's claimed aggregation process.

Second, the Examiner cites Snyder, Col. 16, lines 52-58. This section describes the rendering of a page layout including graphics. Again, this section does not address the aggregation of documents. Applicant submits that this section is not relevant to applicant's claimed aggregation process.

Third, the Examiner cites Snyder, Col. 20, lines 20-24. This section describes the storage of HTML document in a buffer. The complete paragraph in Snyder states that:

In view of the communication delays and to maintain the pace, it is presently preferred that 32 web agents of type A operate in conduction with each web agent of type B. Thus a plurality of web agents of type A continuously fetch and feed into a buffer or queue all web page files of targeted web pages, including their source code and their graphic images, such as JPG, GIF, Java, Flash, etc., all being stored locally. One or more web agents of type B, preferably one for a number of Agents A (e.g., 32) continuously processes and removes files from this buffer to produce and render one web page snapshot image after another. Concurrently with this process, the text portion of the web page data is indexed or categorized. (Snyder, Col. 20, lines 11-24, emphasis added)

Again, this section cited by the Examiner is silent regarding a referencing document that references a referenced document or an aggregation of documents. What this section of Snyder does disclose is that the entire content of a web page is fed into the buffer and concurrently, the text portion of the web page data is indexed or categorized. Thus, this section of Snyder is irrelevant because it does not disclose any aggregation of documents prior to indexing. Indeed, for a given web page given as input, Snyder discloses that "the text portion of the web page data is indexed or categorized," and this is performed without any aggregation being involved whatsoever.

Fourth, the Examiner cites Snyder, Col. 24, lines 17-20. This section describes a crawling step carried out by an Agent "A." The complete paragraph in Snyder states that:

In queuing embedded links found on pages for processing, Web Agent A distinguishes framed links from direct links. When processing a framed page, preferably, the crawler invokes the framed page's internal links to find and queue additional links, but does not treat every framed link as a new web page. Insofar as Web Agent A encounters websites with frames, it processes the data local to that web site and checks for the presence of a website with frames. When a frame page is detected, the Web Agent A checks for a download complete message (end-of-file) for every framed element and processes the text and graphics of the frame and the contents both. (Snyder, Col. 24, lines 10-20, emphasis added).

Based on the above four sections in Snyder, the Examiner concludes that 1) Snyder indexes the textual content of a webpage and 2) that the textual content of a webpage is generated by taking the content of the framing page (referencing document) and the content of the frame (referenced document). According to the Examiner's reading of Snyder, Snyder would provide the same index as applicant's claimed invention as it indexes the content of each webpage (purported *aggregate* document), which would include the main page and the referenced pages for each web page. Applicant indicated in its previous response that this was an incorrect reading of Snyder.

In Snyder, when first processing a web page which contains frames, such frames have already been defined in the main page (see Snyder, Col. 24, line 18, above "When processing a framed page..."). In Snyder, the web page processed by the crawling agent is already a document comprising one or more frames. When a web page with frames is detected, *internal links* are used by the crawler such that the whole content of the web page (including that of internal frames) are stored in the buffer. Snyder does NOT aggregate an internal frame to its corresponding main page because Snyder first begins processing a web page comprising BOTH the main and the internal frames. In Snyder, the content of both the main and the internal frames

are present at the beginning of processing, thus there is no aggregation as is recited in applicant's independent claims.

Once the whole content of the web page is stored in Snyder, the whole content is indexed in relation to the URL of the framing page. In that regard, Snyder states that:

The database 62 contains an index developed from automatic analysis (generally "indexing") or human review (categorization) of the text and other data, indexed to the URLs of the pages from which they were obtained (Snyder, Col. 9, line 65 - Col. 10, line 2, emphasis added).

Accordingly, the initial content of a web page is indexed in relation to the URL of that web page. For one web page as input (e.g., one document), Snyder recovers index terms related to one URL. Given the above, it is clear that Snyder does not aggregate two documents, wherein one document references another document, nor does Snyder teach or suggest that an index has been generated for that purported *aggregated* document. Further, in Snyder, only one document (a web page) is considered for indexing.

Snyder Teaches Away

Applicant submits that Snyder teaches away from applicant's claimed invention, in many respects. For example, if one considers that frames are independent HTML documents, what would happen if one attempts to *aggregate* documents in view of the Examiner's conclusions based on Snyder? In such a scenario, for a given web page as input to the Snyder process, such a web page would be analyzed and processed such that the entire content (including that of internal frames) is indexed to the URL of the main page. Accordingly, if ten web pages are given as inputs to the Snyder process, ten pages would be indexed in relation to the ten URLs, with each of the URL's corresponding to respective main pages. (See Snyder, Col. 9, line 65 to Col. 10, line 2, "indexed to the URLs of the pages from which they were obtained"). Thus, in this

example according to Snyder, the final outcome results in as many documents as there are web pages in the input.

Such combined documents are not aggregated documents, and the processing described in Snyder is clearly not the same as aggregating the documents as recited in applicant's claimed invention. In contrast, in applicant's claimed invention, processing first aggregates a referencing document and a referenced document, wherein the referenced document is referenced in the referencing document by a link. Accordingly, if ten web pages are given as inputs, wherein for example, five of the web pages point to the remaining five pages, at most only five truly aggregated documents are built and then indexed (and not in relation to initial URLs but in relation to the resulting aggregated documents). Accordingly, there are fewer aggregated documents than web pages processed in input, which is very different than the Snyder process.

Further, Snyder requires that index terms be related to each link or URL from which they were obtained by relying on a classical indexer (Col. 9, lines 61-67). In Snyder, there is no reduction in the number of documents. In contrast, in applicant's claimed invention, fewer aggregated documents are produced compared with number of web pages taken as inputs. This is supported in applicant's specification, which states that:

In the example of figures 1 and 2, the index would consider that both frames 2 and 4 as well as the various documents referenced in these frames form a single document. In other words, documents 30, 32, 34, 36, 38, 40 and 42 would be considered as a single aggregate or logical document. In the table of aggregate documents, the single logical documents would be associated with physical documents 30, 32, 34, 36, 38, 40 and 42. (Specification, page 8, 4th par., emphasis added).

Accordingly, in an embodiment of applicant's invention, external HTML documents (such as documents 36) referenced in a frame of another HTML document are aggregated together for a subsequent combined indexing. This result cannot be obtained from Snyder,

because in Snyder, the web agent, which distinguishes between internal links and direct links (e.g. to other documents as *the various documents referenced in these frames*), would restrict the purported *aggregation* of content to that of the frames included in an HTML document, and would not include the content of *the various documents referenced in these frames*. Because the content combined in the output of Snyder corresponds exactly to the initial content of the web page that contains frames, no aggregation occurs.

In summary, with respect to aggregation and indexing, in Snyder, the content of both the main page and the internal frames are present from the beginning of processing in a single document (the main web page), thus there is no aggregation as recited in applicant's claimed invention. Rather, Snyder should be said to begin with an already combined document (namely, the web pages containing frames) and simply indexes the entire content of the combined document. This does not involve an aggregated document not existing as a document in said collection of documents, as claimed.

Links or URL

Further, claim 1 has been amended to recite "a referencing document and a referenced document referenced in the referencing document by one or more links". As disclosed in Snyder and in the tutorial mentioned by the Examiner on page 6 of the Office Action (www.W3schools.com/html/html_frames.asp), frames are defined in the main page by means of internal links, which are not comparable to links, such as Universal Resource locators, inasmuch as internal links lead a browser to reconstruct a full universal locator. For example: `ref = <"http://www.microsoft.com/">` is a link to a page on the World Wide Web, which differs from local internal links corresponding to frames. For instance, in the example below, a frameset with two columns is defined. The first column is set to 25% of the width of the browser window,

while the second column is set to 75% of the width of the browser window. The HTML document "frame_a.htm" is put into the first column, and the HTML document "frame_b.htm" is put into the second column:

```
<frameset cols="25%,75%">  
<frame src="frame_a.htm">  
<frame src="frame_b.htm">  
</frameset>.
```

Such a definition of frames cannot be understood as being independent from the locator of the main page wherein where frames are defined. Therefore, the way the frames are defined in the main page makes use of local internal links, which markedly differ from universal resource locators. Processing of on such internal links is contrary to applicant's claimed invention because it cannot allow for providing an aggregated document that does not exist as a document in the collection of documents. On the contrary, making use of local internal links indicates starting from an already existing document (the main page). Accordingly, Snyder does not teach or suggest aggregating documents so as to achieve an aggregated document not existing as a document in the initial collection of documents, but instead Snyder teaches recomposing components (such as frames) of an already existing document.

Accordingly, applicant submits that independent claim 1 is not anticipated by Snyder and is allowable. Applicant further submits that dependent claims 10 are 14 are allowable as depending from an allowable base claim.

Rejection under 35 U.S.C. §103

Claims 8 and 9 stand rejected under 35 U.S.C. §103 as being unpatentable over Snyder in light of Bourdoncle (2002/005894). Applicant respectfully traverses the rejection. Applicant reasserts that above arguments made under §102 to traverse the rejection under §103. Snyder is completely deficient because it is missing several important elements of applicant's claimed

invention, such as the process of aggregating the documents and indexing the document after aggregation. Bourdoncle is similarly deficient because it is missing the same elements and does not provide, either by teaching or suggestion, any of the missing elements. Accordingly, applicant submits that the combination of Snyder and Bourdoncle does not, and cannot teach or suggest applicant's claimed aggregation process.

Closing Remarks

Pending claims 1 and 8-14 are believed to be patentable. Applicant respectfully requests the Examiner grant early allowance of this application. The Examiner is invited to contact the undersigned attorneys for the applicant via telephone if such communication would expedite this application.

Respectfully submitted,



Registration No. 38,110
Eric D. Cohen
Attorney for Applicant

BRINKS HOFER GILSON & LIONE
P.O. BOX 10395
CHICAGO, ILLINOIS 60610
(312) 321-4200